Support for Maui Space Surveillance Site and Maui High Performance Computing Center

February 1999

Prepared by:

Schafer Corporation
2000 Randolph Rd. S. E.
Suite 205
Albuquerque, NM 87106

Task Report - Naval Research Laboratory
Contract N00014-97-D-2014/001

Support for Maui Space Surveillance Site and Maui High Performance Computing Center

Schafer Corporation

2000 Randolph Rd., SE, Ste. 205

Albuquerque, New Mexico 87106

The work described in this document was accomplished by Mr David Tyler (Schafer) and Ms Kathy Schulze, a Schafer consultant for advanced and parallel computer programming. All work presented here was funded either by the Maui High Performance Computing Center (MHPCC) Research and Development Consortium under a cooperative agreement between the Air Force Research Laboratory (AFRL) and the University of New Mexico or through AFRL/DEBI. Attached is a paper describing work performed by Schafer under this subtask (Tyler and Stribling). This paper was presented at the ESO/OSA Conference on Astronomy with Adaptive Optics. Although the paper was presented after the lead author left Schafer, the paper acknowledges that the work was performed while the author was employed at Schafer. Also attached is a paper describing work on analytical models of constraint-induced noise and information transport in imagery degraded by atmospheric turbulence (Prasad and Tyler). This paper was invited and presented at the UNM Center for Advanced Studies Workshop on Fundamental Issues in Image Formation, Detection, and Processing.

Summary: Integration and test of the GEMINI instrument package. Most of the work performed under this task supported development of the GEneralized Multi-wavelength INfrared Imager (GEMINI) for the Maui Space Surveillance Site. GEMINI, not to be confused with the National Science Foundation's Gemini Telescopes Project, is a one-of-a-kind sensor package built for USAF Space Command operational use in conjunction with its space object identification (SOI) mission. GEMINI hosts a variety of sensors for use with several algorithms for reconstructing high-resolution satellite and astronomical images from data degraded by atmospheric turbulence.

The first GEMINI operating mode made available to Space Command delivers high-resolution images using a speckle imaging technique. Schafer contributed to the project by implementing and developing parallel speckle imaging recovery and calibration algorithms and by designing computer simulations to model real-world detector effects and predict available resolution. Schafer also conducted an observational evaluation of the speckle imaging algorithm.

A distinctive feature of the GEMINI data acquisition/reduction system is the provision for near real-time (NRT) video feedback of reconstructed imagery to the GEMINI operator. NRT feedback allows the operator to select camera exposure times and optical filters based on the visual quality of reconstructed data. These operating parameters can thus be optimized *while data is being acquired*, potentially eliminating the need for subsequent observations of the same object. For example, raw image data of a bright satellite can yield a poor reconstruction if most of the light is concentrated in a strong, localized glint. Selection of a different filter may reduce the glint effect in the reconstruction, which would be apparent to the operator using an NRT imaging system. Also, NRT reconstruction can be used to reduce even large amounts of data is a timely manner, allowing fast turnaround between observing and imaging without simply ignoring most of the data (as was commonly done at the MSSS prior to GEMINI).

The speed required for NRT imaging was obtained with a parallel implementation of speckle imaging reconstruction and calibration algorithms. The code is executed on the MHPCC's IBM SP2 supercomputer. Schafer was responsible for design, parallelization, and optimization of all image reconstruction and calibration software, as well as joint development of the entire data acquisition system with a multi-contractor team.

During the period March-Nov 1998, Schafer participated in laboratory and operational testing of GEMINI hardware and softare. The data acquisition and processing pipeline was a multi-contractor team effort, with Schafer responsible for parallel (NRT) execution of the speckle imaging algorithm. This software, along with Schafer-developed calibration algorithms, was tested with lab data and telescope data, once GEMINI was installed on the MSSS 1.6-m telescope. Schafer was also responsible for evaluating GEMINI resolution under a variety of imaging conditions. Sample "first light" results, of a binary star and the orbiting Hubble Space Telescope, are shown below. Figure 1 shows a surface plot of uncalibrated ("raw") data from an observation of the binary star ADS 10598. Note that due to the corrupting effects of the Earth's atmosphere, the intensity distribution bears no discernable resemblance to that of a binary star. In Fig. 2, where the result of processing the data with the Schafer implentation of the speckle imaging algorithm is shown, the two components of the binary are clearly visible and resolution is nearly diffraction-limited. Figure 3 shows a processed image of the Hubble Space Telescope. Note the aperture cover is clearly visible in the "open" position, and the slight "twist" in the Hubble's solar panels, seen in photos taken by space shuttle crews, is readily apparent.

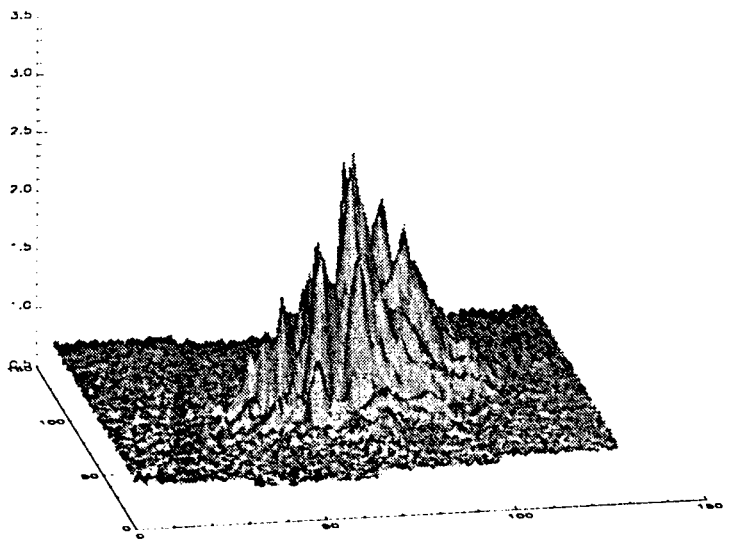

Fig. 1: Uncalibrated frame from ADS 10598 data ensemble.

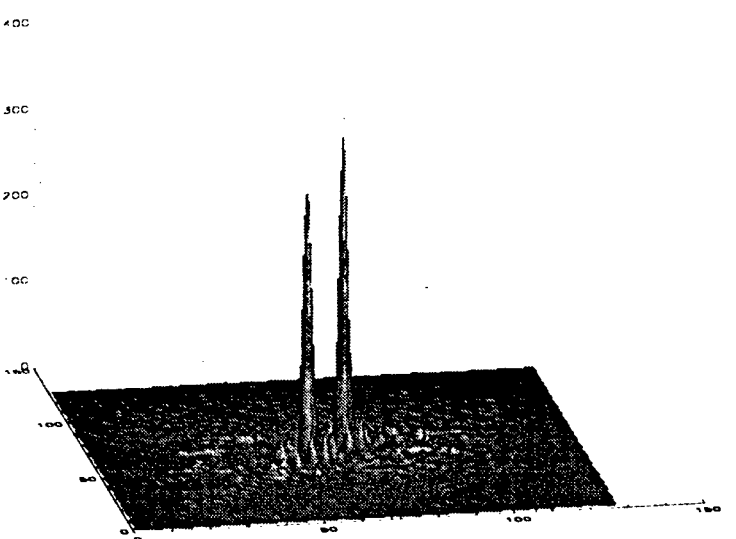

Fig. 2: Bispectrum speckle imaging recovery of ADS10598.

Fig. 3: GEMINI image of the Hubble Space Telescope

Summary. Noise and information transport in imagery. Previous work by Schafer has demonstrated a practical noise-removal scheme utilizing projections onto convex sets. During March-November 1998. Schafer developed a code to calculate the information content of imagery and evaluate the effect of the convex projections scheme on the two-dimensional information distribution of processed image data. The code was used to validate an analytical treatment of the noise and information "transport" phenomenon previously observed. Details are in the attached paper by Prasad and Tyler.

Summary. A wavefront sensor-driven variable geometry pupil for ground-based aperture synthesis imaging. Schafer developed the concept of a variable geometry pupil to increase image resolution for ground-based near-IR and optical telescope imaging. In this scheme. a curvature-type wavefront sensor provides an estimate of the wavefront curvature to

the controller of a high-resolution spatial light modulator (SLM) or micro-electromechanical (MEM) mirror, positioned at an image of the telescope pupil. This optical element, the VGP, passes or reflects the incident beam only where the wavefront phase is sufficiently smooth, *viz.*, where the curvature is sufficiently low. Using a computer simulation, Schafer showed the VGP can sharpen and smooth the long-exposure PSF and increase the OTF SNR for tilt-only and low-order AO systems, allowing higher resolution and more stable deconvolution with dimmer AO guidestars. Details are in the attached paper by Tyler and Stribling.

A wavefront sensor-driven variable geometry pupil for ground-based aperture synthesis imaging

D.W. Tyler[a] and B.E. Stribling[b]

[a] University of New Mexico, Albuquerque High Performance Computing Center
1601 Central Ave. NE, Albuquerque NM 87131 USA

[b] Michigan Technological University, Dept. of Electrical Engineering
1400 Townsend Dr, Houghton MI 49931 USA

ABSTRACT

We describe a variable-geometry pupil (VGP) to increase image resolution for ground-based near-IR and optical imaging. In our scheme, a curvature-type wavefront sensor provides an estimate of the wavefront curvature to the controller of a high-resolution spatial light modulator (SLM) or micro-electromechanical (MEM) mirror, positioned at an image of the telescope pupil. This optical element, the VGP, passes or reflects the incident beam only where the wavefront phase is sufficiently smooth, *viz.*, where the curvature is sufficiently low. Using a computer simulation, we show the VGP can sharpen and smooth the long-exposure PSF and increase the OTF SNR for tilt-only and low-order AO systems, allowing higher resolution and more stable deconvolution with dimmer AO guidestars.

Keywords: Adaptive optics, Aperture synthesis, Optical instrumentation

1. INTRODUCTION

Time-varying, spatially random refractive index fluctuations in Earth's atmosphere limit the resolution of even the largest ground-based telescopes to that obtainable by a telescope of diameter equal to the Fried seeing parameter, r_o. This limitation can be understood by considering the effects of *pupil redundancy* in imaging while viewing image formation itself as an interferometric process.[1] In the interferometric model, the pupil of an Earth-based telescope is seen as an ensemble of r_o-size "apertures" or patches, each pair of which forms an interference pattern on the detector. The fringe patterns have a variety of orientations and modulation frequencies depending on the orientation and separation of these patches, and their incoherent sum is the detected image. In an unobscured telescope pupil in the presence of atmospheric turbulence, there may be many r_o patches separated by identical distances and having the same orientation. This means the fringe pattern in a given direction on the detector with a given frequency is actually the sum of many individual *redundant* fringe patterns. Due to turbulence, the phase difference between any two redundant fringe patterns is a random variable, reducing the visibility of the sum fringe pattern and limiting the resolution of the detected image. Use of adaptive optics (AO) increases resolution by reducing the relative shift of the individual fringe patterns and increasing the spatial coherence of the field incident at the telescope pupil, thus increasing the visibility of the sum pattern.[2] Another, less complicated approach is to limit redundancy by masking selected portions of the telescope pupil. Several relatively small openings can be used to synthesize a large, high-resolution telescope. While the phases of the resulting fringe patterns are still randomized by turbulence, the resulting increase in visibility *amplitude* SNR can be used to enhance phase recovery by such methods such as speckle imaging or self-calibration.[3] Of course, masking the light-gathering area of the telescope decreases the available signal and passes a reduced set of spatial frequencies. This technique is known variously as "optical aperture synthesis," "pupil" or "aperture masking," "non-redundant masking", and "masked-aperture speckle imaging." We use the nomenclature "partially-redundant masking" (PRM) as a general description of the several variations on the above theme.

Since the first astronomical images reconstructed using aperture masking[4] were reported in 1988, considerable effort has been expended by the astronomical community to quantify the viability of this technique,[5–10] especially in comparison with adaptive optics systems.[11] Initial studies and experiments were done using masks with a fixed (small) number of r_o-size openings. To obtain reasonable u-v coverage with a small number of apertures, researchers applied several rotations of the mask relative to the incident beam. More recently, workers employed more novel geometries, such as an r_o-size annulus[5,12] or a slit.[13] These geometries are not actually "non-redundant," hence the use of "PRM." Generally, AO is considered to be the more robust of the two methods, allowing deeper imaging at

higher resolution. As a consequence, there has been some interest in combining the techniques.[14] In this paper, we describe an adaptive pupil mask for use with AO, driven by the AO wavefront sensor. As we will show, the concept has potential to augment the resolution provided by conventional AO.

The variable-geometry pupil (VGP) we propose here is a new type of adaptive optical element, correcting the amplitude of complex visibilities rather than the phase, and realizing PRM concepts in an adaptive, controllable way. Such an element might be a high-resolution spatial light modulator (SLM) or a micro-electromechanical (MEM) mirror. The VGP would be placed between the AO deformable mirror and the science instrument and controlled in such a way that only portions of the incident beam with sufficiently small phase variation would be passed (SLM) or reflected (MEM) to the science instrument. Since we envision the VGP would be used with AO, an estimate of the wavefront phase is readily available from the AO wavefront sensor (WFS). In fact, if the AO system uses a curvature-sensing WFS, the WFS provides the information on phase "smoothness" directly to the VGP controller. In our scheme, the WFS would drive the AO system in a closed control loop while providing "go-to" curvature signals to the VGP: that is, there would be no feedback between the VGP and the WFS. The VGP would thus receive commands at intervals on the order of the inverse AO bandwidth, passing or reflecting light in response to the random atmospheric phase perturbations and the AO system corrections. Since the phase of the wavefront incident at the WFS and VGP (modulo noncommon optical path errors) is a random process, the shape of the VGP would also change randomly, potentially providing full u-v coverage after a few WFS commands. The VGP, as we envision it, would overcome many of the limitations in the fixed-pupil PRM concept; for example, the VGP can be adjusted for seeing and object brightness conditions without modifying or replacing the mask, and there is no need for beam or mask rotation. We note that the concept of selectively masking the telescope pupil during AO operation is also discussed in Morossi, *et al.*[15] in these proceedings.

In the following section, we describe the computer simulation methods used to model the telescope, AO, and VGP. In Section 3, we present simulation results, showing PSFs and OTF SNR plots with and without the VGP for various AO systems. We summarize our work in Section 4.

2. COMPUTER SIMULATION AND MODELING

We model incoherent imaging with a 3.6-m telescope and two AO systems. The computer simulation we use generates independent random realizations of Kolmogorov turbulence. The AO subroutines actually model a Shack-Hartmann WFS rather than a curvature sensor, but we calculate curvature separately after generating the post-AO phase map. For the simulation results shown here, both the AO and VGP are driven by "go-to" commands in open loop.

After AO compensation, the wavefront curvature is calculated from

$$C(\mathbf{r}) = \mathcal{F}^{-1}\{\rho^2\Phi(\rho)\}, \tag{1}$$

where $C(\mathbf{r})$ is the curvature map at pupil coordinate $\mathbf{r}$. $\Phi(\rho)$ is the Fourier transform of the pupil phase map $\phi(\mathbf{r})$, and $\rho = |\boldsymbol{\rho}|$. Since the above operations are done on the same grid as the array representation of the $\phi(\mathbf{r})$, $C(\mathbf{r})$ varies on an unrealistically small scale. To model the finite sampling resolution of an actual curvature WFS, we calculate $\hat{C}$ from C by smoothing with a Gaussian filter having a FWHM corresponding to the subaperture size used in the sim. The VGP pupil function is formed by setting the telescope pupil function to zero wherever $\hat{C}$ exceeds a threshold value. A VGP pupil function, typical for the conditions described in the next section, is shown in Fig. 1. Subsequent to the calculation of the VGP pupil function, a generalized pupil function is formed by multiplying the phase screen and VGP pupil function. PSF images are formed from the squared modulus of the Fourier transformed generalized pupil function.

3. SIMULATION RESULTS

Our baseline imaging model consists of a 3.6-m telescope and a 37-actuator AO system. This configuration yields an actuator separation of 55 cm. The Shack-Hartmann WFS model has 30, 48-cm subapertures. In all our simulations, we use a WFS phase measurement standard deviation of 0.18 radians, corresponding to a guidestar magnitude of about 6 with a sampling rate of 1000 Hz. This small error means WFS noise is not the limiting factor in image resolution, although we choose seeing conditions of $r_o = 20$ cm at our imaging wavelength of 750 nm that are somewhat demanding for the AO system ($D/r_o = 18$). To avoid excessive WFS aliasing, we model wavefront sensing at H band, where the effective seeing parameter is ≈ 50 cm. The detector pixels are critically sampling, with each

pixel subtending about 26 milli-arcsec. In these seeing conditions, the AO performance is partially-compensating due to fitting error, allowing for some resolution gain using the VGP, but not unrealistically degraded.

Fig. 1: Sample VGP pupil function. White indicates low-curvature regions. Note the telescope secondary obscuration is not readily apparent because it is contiguous with two regions of high wavefront curvature, which are masked by this realization of the VGP pupil function.

VGP operation is illustrated in Fig. 2. The top part shows a surface plot of the DM-only generalized pupil function. Notice the vertical scaling, indicating the wavefront phase, extends from -6 to +6 radians. While the DM-only AO has corrected much of the phase from the $D/r_o = 18$ seeing, the residual is significant. In the bottom of Fig. 2, we show a surface plot of the VGP/DM generalized pupil function. Note the scaling now extends from -4 to + 4, and regions with large phase excursions in the top part of Fig. 2 have been blocked by the VGP.

In Fig. 3, we show long-exposure log PSFs using the 30-subaperture AO system with and without the VGP. The PSFs were formed by averaging the images from 300 independent turbulence realizations. No detector noise is modeled for these plots, since we desire to see potential PSF improvements prior to noise modeling. To distinguish between the two types of AO, we refer to conventional adaptive optics without the VGP as "DM-only" AO. The use of both the DM and VGP is called "VGP/DM AO." Note the VGP/DM PSF has more energy in the diffraction-limited core and smaller halo features.

If we improve the DM correction by adding more subapertures and actuators, the improvement using the VGP is not so pronounced, as we expected. In Fig. 4, we show the same PSFs for AO with 40 subapertures and 45 actuators. Although the DM-only PSF has improved considerably, the VGP/DM PSF still has more core energy and smaller sidebands. Conversely, degrading the AO performance causes a dramatic increase in the relative performance of VGP/DM AO over DM-only AO. Fig. 5 shows PSFs for a 30-subaperture WFS driving a fast-steering mirror (FSM) for tilt removal and a VGP/FSM combination. In this case, the improvement in core energy using the VGP is nearly an order of magnitude: in fact, there is no discernible core without the VGP.

The implications of these improvements are illustrated in Fig. 6: Here, we have plotted the DM-only PSFs for the 30 and 40-subaperture systems along with the VGP PSF for the 30-subaperture system. Note the VGP PSF with the 30-subaperture AO has nearly the same core energy and lower sidelobes than the 40-subaperture DM-only PSF. Potentially, then, the VGP can extend the resolution obtained with DM-only AO to the dimmer guidestars allowed with fewer, larger subapertures.

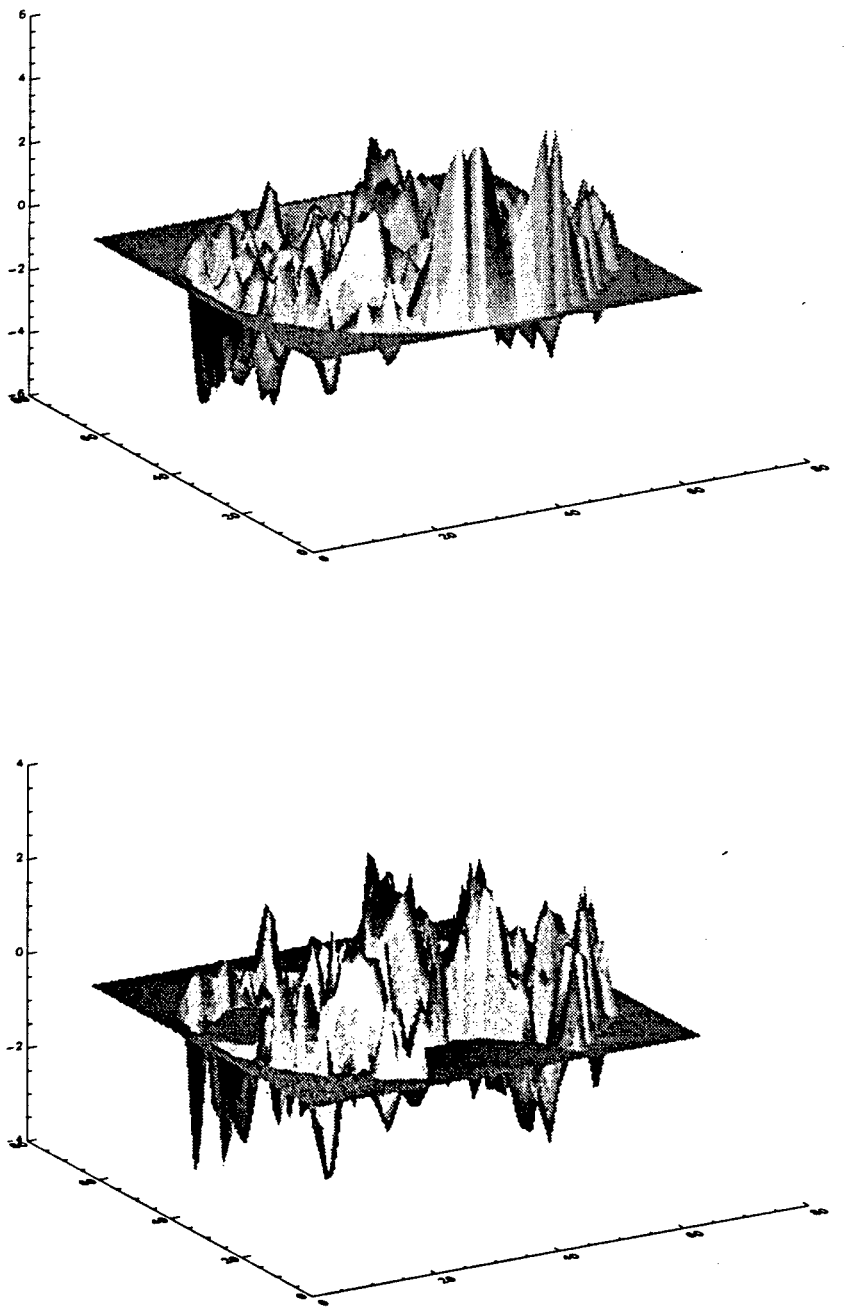

Fig. 2: DM-only (top) and VGP/DM (bottom) generalized pupil functions.

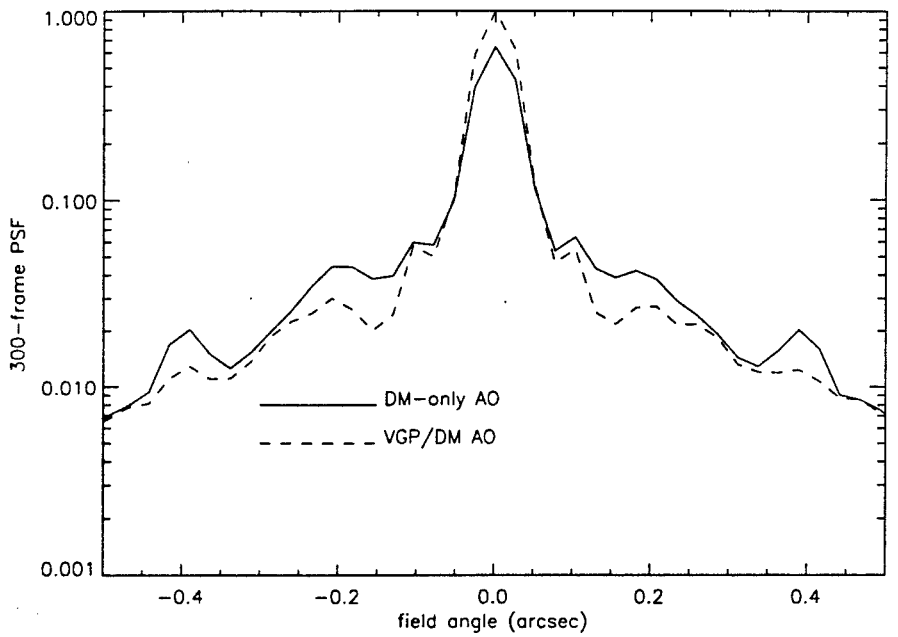

Fig. 3: Long-exposure log PSF cross-sections for a 30-subaperture WFS driving a 37-actuator DM only (solid line) and a VGP/DM combination (dashed line).

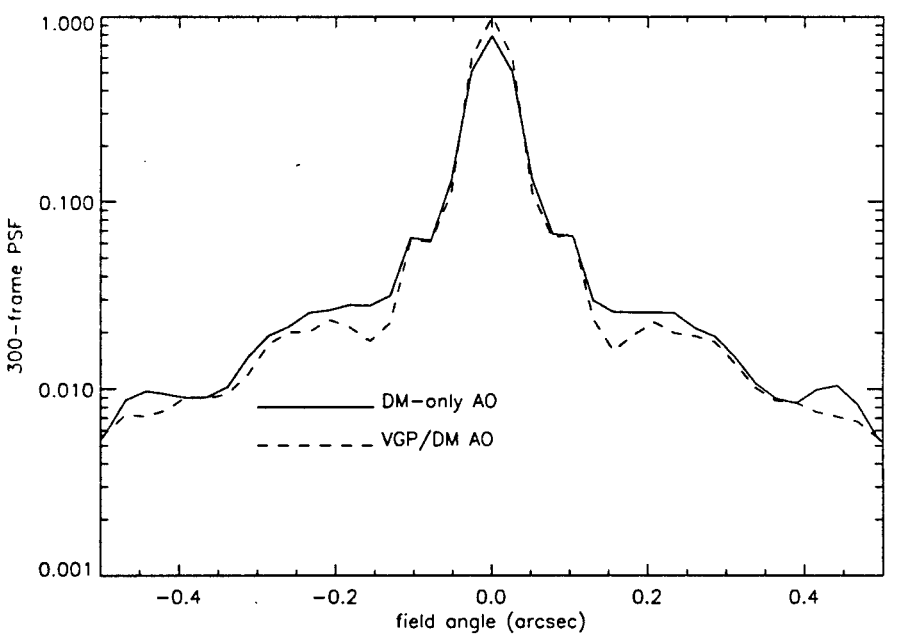

Fig. 4: Long-exposure log PSF cross-sections for a 40-subaperture WFS driving a 45-actuator DM only (solid line) and a VGP/DM combination (dashed line).

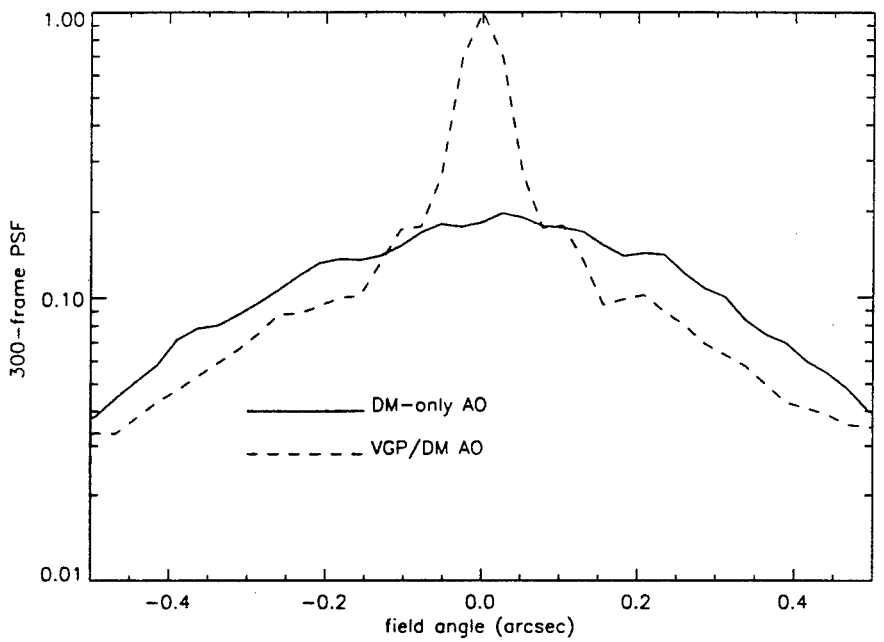

Fig. 5: Long-exposure log PSFs for a 30-subaperture WFS driving a tip-tilt FSM only (solid line) and a VGP/FSM combination (dashed line).

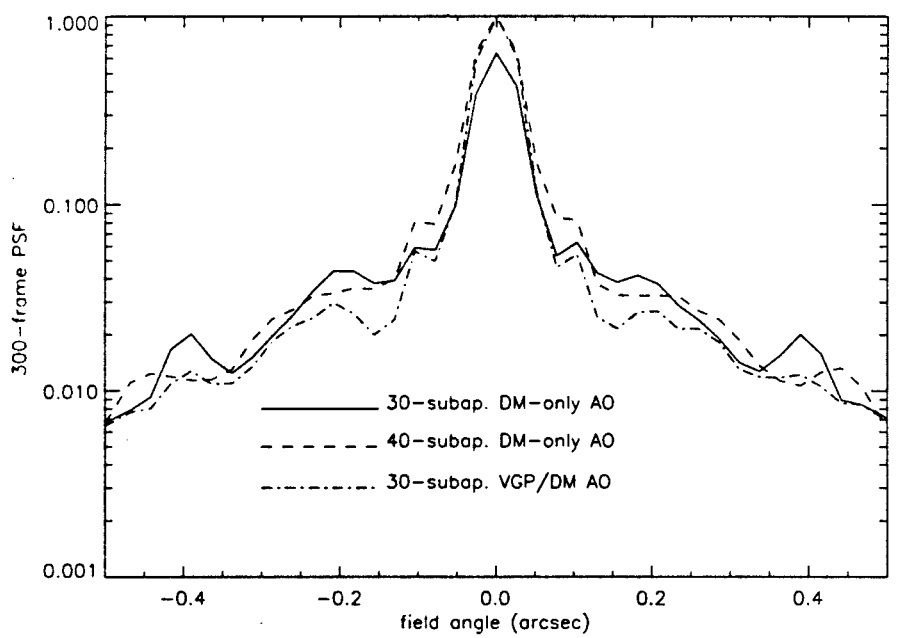

Fig. 6: Long-exposure log PSF cross-sections for a 30-subaperture WFS driving a 37-actuator DM only (solid line). a 40-subaperture WFS driving a 45-actuator DM only (dashed line), and a 30-subaperture WFS driving VGP/DM combination (dot-dashed line).

While improved PSF morphology is desirable, the key to successful deconvolution stability is the OTF SNR. To assess the relative SNR performance of VGP/DM and DM-only AO, we included detector noise in the image formation process and calculated sample SNRs over a 300-frame ensemble. For these simulations, we included Poisson (signal-dependent) noise from a double-precision random number generator and zero-mean Gaussian (additive) "read" noise with an RMS amplitude of 7 ϵ^- per pixel per read.

In Fig. 7, we show the OTF SNR for a magnitude 10 object. The signal level to the WFS is unchanged, since we desire only to study the effect on the OTF SNR of reducing the signal level with the VGP. The 2-d SNR map has been radially averaged to produce a simpler x-y plot. Note the VGP/DM SNR is significantly higher than the DM-only SNR for all frequencies beyond the seeing spike. The relative reduction in VGP/DM SNR near DC is due to the increased fluctuation in total intensity as the VGP aperture area changes randomly; the only total count variance for DM-only AO is due to detector noise.

In Fig. 8, we show the same plots for a magnitude 12 object. The curvature threshold is the same as was used for all other plots shown here. Here, the DM-only SNR curve is slightly higher than the VGP/DM SNR at all frequencies of interest. Clearly, the SNR gains using the VGP are magnitude-dependent. However, we point out that with *fixed* aperture masks, there are two free parameters available to optimize measured visibility SNR[16]; namely, the size and number of the mask openings ("subapertures"). While there are not generally discrete subapertures in the VGP as there are with a fixed aperture mask, the ratio of turbulence noise power to detector noise power can still be manipulated by adjusting the wavefront curvature threshold and by restricting the regions over which the VGP can allow light to pass. For dim objects, larger subapertures are favored in fixed aperture masks, even though the turbulence power in the data increases. By analogy, dim objects should also drive a VGP controller to use a lower curvature threshold, increasing the pupil area. Dim objects also dictate a reduction in the number of fixed subapertures: this reduces the number of baselines among which light must be shared (naturally, this reduction in the number of subapertures becomes academic for objects so dim as to indicate aperture masking should not be used at all). The same effect can be obtained with the VGP by allowing the VGP controller to mask regions of the pupil regardless of the wavefront curvature. For example, the VGP could mask all but a small annulus of the pupil, only opening for low-curvature wavefronts within the area of the annulus. These ideas and optimal control rules will be studied in future work.

4. SUMMARY

We have described a variable-geometry pupil (VGP) to increase image resolution for ground-based near-IR and optical imaging. A curvature-type wavefront sensor provides an estimate of the wavefront curvature to the controller of a high-resolution spatial light modulator (SLM) or micro-electromechanical (MEM) mirror, which passes or reflects the incident beam only where the wavefront phase is sufficiently smooth, *viz.*, where the curvature is sufficiently low. Using a computer simulation, we have shown the VGP can sharpen and smooth the long-exposure PSF and increase the OTF SNR for tilt-only and low-order AO systems, allowing higher resolution and more stable deconvolution with dimmer AO guidestars.

We point out that the VGP can be used for long-exposure imaging, in contrast with fixed-aperture masks, which are typically used to enhance visibility SNRs in speckle imaging schemes. Also, we note that while the VGP/DM PSFs shown here have more core energy and smaller halo features close to the core, there is generally more energy in the VGP/DM halo at radii greater than about an arcsecond from the core. This is due to the fact that the VGP aperture area is bounded from above by the DM-only aperture area, and is most always smaller. Additionally, we note that where a curvature-sensing WFS is used, the VGP uses curvature information directly, rather than after integration, meaning the VGP controller is never subject to the numerical errors that can occur using a large number of subapertures in a curvature sensor. Finally, we point out that the sharper core obtained with the VGP could increase the sky coverage of an AO-compensated spectrometer[17] or the sensitivity of a coronographic imager.

ACKNOWLEDGMENTS

The authors gratefully acknowledge support from the USAF/MHPCC R&D Consortium. DWT acknowledges support from the Air Force Office of Scientific Research. When the work described in this publication was performed, DWT was employed at Schafer Corporation, Albuquerque, NM: he is personally grateful to Dr Rettig Benedict of Schafer and many others for his great experiences there. At the same time, BES was employed at the Air Maui Optical Station, AFRL/DEBI, Kihei, Maui, HI.

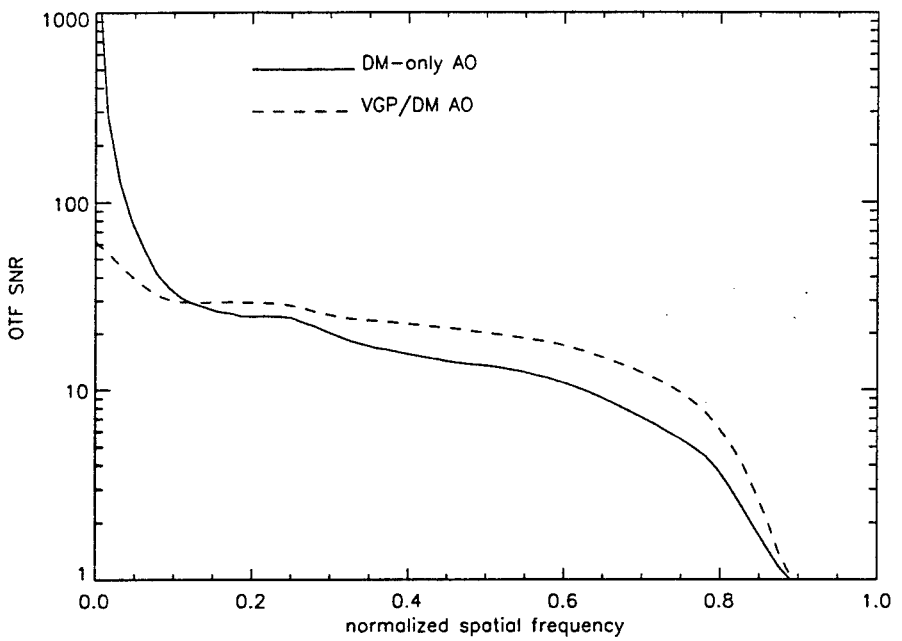

Fig. 7: OTF SNR curves for DM-only (solid line) and VGP/DM (dashed line) AO for a magnitude 10 object.

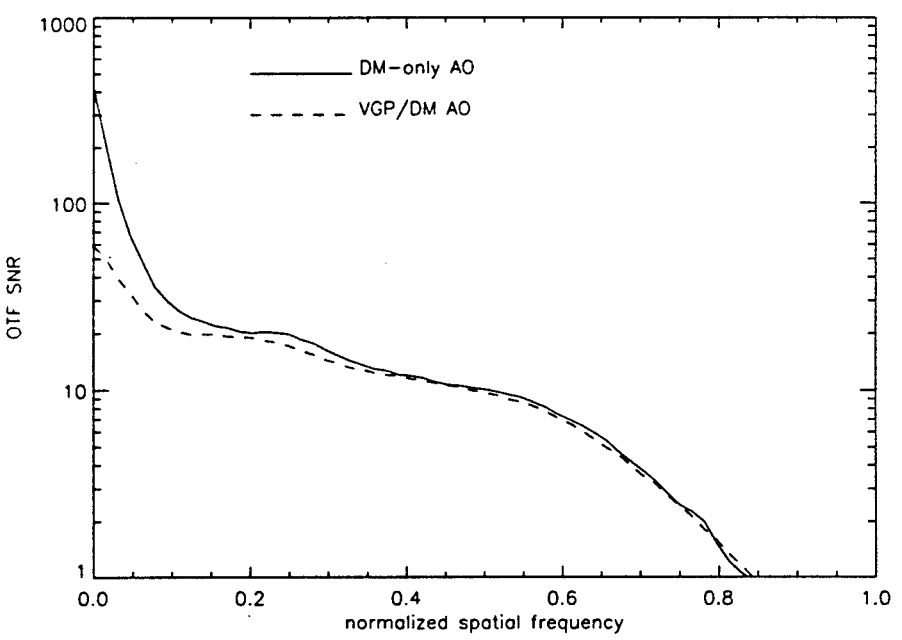

Fig. 7: OTF SNR curves for DM-only (solid line) and VGP/DM (dashed line) AO for a magnitude 12 object.

REFERENCES

1. J.W. Goodman, "Image Formation as an Interferometric Process", in *Statistical Optics*, p. 331, John Wiley and Sons, New York (1985)
2. M.C. Roggemann, *et al.*, "Compensated speckle imaging: theory and experimental results," Appl. Opt. **33**, p. 3099 (1994)
3. T.J. Cornwell, "Radio interferometric imaging of weak objects in conditions of poor phase stability: The relation between speckle masking and phase-closure methods", Astron. Astrophys. **180**, p. 269 (1987)
4. C. Haniff, *et al.*, "The first images from optical aperture synthesis", Nature (London) **328**, p. 694 (1988)
5. C.A. Haniff, D.F. Buscher, J.C. Christou, and S.T. Ridgway, "Synthetic aperture imaging at infrared wavelengths," Mon. Not. R. Astr. Soc. **241**, pp. 51p-56p (1989)
6. T. Nakajima, *et al.*, "Diffraction-limited imaging II: Optical aperture synthesis imaging of two binary stars", Astron. J. **97**, p. 1510 (1989)
7. J. Robertson, *et al.*, "High resolution imaging by optical aperture synthesis — first results from the MAPPIT project", Proc. Astron. Soc. **9**, p. 162 (1991)
8. W.N. Weir, *et al.*, "Infrared non-redundant mask imaging at Palomar", in *Amplitude and Intensity Spatial Interferometry*, J. Breckinridge, ed., Proc. SPIE **1237**, p. 274 (1990)
9. R.W. Wilson, J.E. Baldwin, D.F. Buscher, and P.J. Warner, "High-resolution imaging of Betelgeuse and Mira," Mon. Not. R. Astr. Soc. **257**, pp. 369-376 (1992)
10. C.A. Haniff, A.M. Ghez, P.W. Gorham, S.R. Kulkarni, K. Matthews, and G. Neugebauer, "Optical aperture synthetic images of the photosphere and molecular atmosphere of Mira," Astr. J. **103**, pp. 662-1667 (1992)
11. T. Nakajima and C.A. Haniff, "Partial adaptive compensation and passive interferometry with large ground-based telescopes", Pub. Astr. Soc. Pacific **105** (1993)
12. C. Haniff and D. Buscher, "Diffraction-limited imaging with partially redundant masks I: Infrared imaging of bright objects", J. Opt. Soc. Am. A **9**, pp. 203-218 (1992)
13. D. Buscher and C. Haniff, "Diffraction-limited imaging with partially redundant masks II: Optical imaging of faint sources", J. Opt. Soc. Am. A **10**, pp. 1882-1894 (1993)
14. C.A. Haniff and R.W. Wilson, "Closure-phase imaging with partial adaptive correction," Pub. Astr. Soc. Pacific **106**, pp. 1003-1014 (1994)
15. C. Morossi, M. Franchini, and S. Furlani, "Angular resolution improvement for an 8-meter class telescope via on-line subaperture selection," these proceedings.
16. D.W. Tyler, "Optimal pupil sampling for ground-based astronomical imaging I: Fixed-geometry pupils," *submitted* to JOSA-A.
17. D.W. Tyler and B.L. Ellerbroek, "Sky coverage calculations for spectrometer slit power coupling with adaptive optics compensation," in *Adaptive Optics Technologies*, Proc. SPIE 3353, D. Bonaccini and R.K. Tyson, *eds.*, pp. 201-209 (1998)

Information Dynamics in Constrained Image Deconvolution

SUDHAKAR PRASAD AND DAVID TYLER

Center for Advanced Studies and
Department of Physics and Astronomy
University of New Mexico
Albuquerque, New Mexico 87131

February 1, 1999

ABSTRACT

When measured data are limited in quality and in quantity as they in general are, the additional information contained in *a priori* constraints can often lead to unique image reconstruction in a variety of algorithms. In a linear deconvolution algorithm devised by Tyler and Matson,[1] an *a priori* knowledge of the support of the object was exploited to reduce the atmospheric noise in the measured astronomical data. Prasad[2] has recently showed that such noise reduction can be described in terms of a drift-diffusion dynamics. In the present paper, we consider the problem of constrained image deconvolution from Shannon's mutual information viewpoint. We shall show that the noise dynamics induces a corresponding information dynamics as well, leading to the possibility of assigning an information-theoretic measure to the usefulness of constraints in image processing.

I. Introduction

Statistical information, as defined by Shannon,[3] refers to the amount of information one *gains* about the input by measuring the output of a communication channel. In the absence of noise or another limitation, the output furnishes a complete specification of the input, and the communication can be thought of as being perfect. By regarding the most general imaging system as a communication channel, with object inputs and image outputs, it is clear that one can characterize the performance of an imaging system too in terms of the information it processes.

The use of information theory to characterize image formation was first discussed by Fellgett and Linfoot[4] not long after Shannon's seminal work. They addressed two competing viewpoints about the quality of an image. The first viewpoint is the obvious one, namely that the image which resembles the object most is the best image. The second viewpoint is the information theoretic one. The best image must carry the most information about the object. That the second viewpoint is generally superior is at once clear when one recognizes that often the object is not known *a priori* so the question of resemblance of the image to the object cannot even be answered in such situations. Fellgett and Linfoot considered some simple examples of object ensembles and additive noise with Gaussian probability distribution in the same spirit as Shannon's Gaussian channels, but further use of information theory to describe imaging had to wait until Frieden's work in the early seventies.

Taking the lead from Jaynes's interpretation[5] of statistical mechanics from the information theoretic viewpoint. Frieden[6] introduced the notion of maximum entropy to characterize the least biased object intensity distribution $O(\vec{x})$ that could have given rise to the measured image data. The maximum entropy method (MEM) has slowly risen in importance to become one of the most commonly used algorithms in modern image processing. A somewhat different method, based on maximizing the mutual information rather than the entropy, was also proposed some years later by Frieden,[7] who compared the two approaches. In both approaches the intensity at pixel $\vec{x}$ is taken to be proportional to a probability function $p(\vec{x})$ and statistical considerations concern a *single* object and a *single* image, rather than statistical ensembles of objects and images.

A priori constraints like positivity and a knowledge of object support can provide further information on the object and constrain the reconstruction of the object from limited data. It has been shown by Tyler and Matson[1] that the imposition of the support constraint not only limits the physical extent of the reconstructed object in a trivial manner but it also enables a reduction of noise inside the support, thereby improving the quality of the measured data. This noise reduction algorithm was analyzed by Prasad[2] who modeled the noise transport in this algorithm as a drift-diffusion process in the spatial frequency plane.

In the present paper, we analyze the dynamics of mutual information throughput of an imaging system, in much the same spirit as Fellgett and Linfoot,[4] without and with the application of an *a priori* knowledge of the object support. The drift-diffusion dynamics of Fourier noise induces an analogous dynamics of mutual information through the imaging system on imposition of the support constraint. This will be the main focus of the present work.

II. Noise Dynamics under Support Constraint

The problem of interest here is an astronomical imaging problem, in which a ground-based telescope forms the image of an object in the sky. With no or incomplete adaptive correction of the wavefront errors caused by atmospheric turbulence, the raw image formed by the telescope is correspondingly noisy and poorly resolved even under the best seeing conditions. The image formation is most simply described in the Fourier (or spatial-frequency) domain, where the telescope and the atmospheric turbulence combine to form a spatial filter with a noisy multiplicative optical transfer function (OTF) $H(\vec{u})$ as a function of the spatial frequency $\vec{u}$. The Fourier transform of the (fluctuating) image intensity is related to that of the object intensity as

$$I(\vec{u}) = O(\vec{u})H(\vec{u}). \tag{2}$$

We assume that there are no other sources of image noise, so the multiplicative noise arising from the OTF H, noise that is evidently non-stationary, is the only one we consider here.

For a circular telescope pupil of large diameter, the wavefront is composed of many statistically independent patches, as is the OTF. Therefore, by the central limit theorem, the statistics of the OTF $H(\vec{u})$ is described approximately by a Gaussian probability density function (pdf), which is also circular about the mean value:

$$P_H(H(\vec{u})) \approx \frac{1}{\pi\sigma_H^2(\vec{u})} e^{-|H(\vec{u})-\langle H(\vec{u})\rangle|^2/\sigma_H^2(\vec{u})}, \tag{3}$$

in which $\sigma_H^2(\vec{u})$ is the variance $\langle |H(\vec{u}) - \langle H(\vec{u})\rangle|^2\rangle$.

The restriction of object support is obviously a restriction of the physical domain. Let $s(\vec{x})$ be the support function, taking the value 1 inside the support and 0 outside. Any intensity that is outside the support can only be noise, which can then be set to 0 enabling thus a trivial reduction of noise. The support-constrained image intensity $i_s(\vec{x})$ at location $\vec{x}$ is thus just the raw image intensity $i(\vec{x})$ at that location times $s(\vec{x})$. In the Fourier domain this becomes a convolution,

$$I_s(\vec{u}) = \int I(\vec{u} - \vec{u}')S(\vec{u}')d\vec{u}'. \tag{4}$$

By subsituting expression (2) for the raw image spectrum* $I(\vec{u})$ into Eq. (4), we see that the noisy OTF $H(\vec{u})$ induces a noise in the support-constrained image spectrum via a convolution.

When uncompensated turbulence-induced errors are significant, we can show[2] that the noise in the real and imaginary parts of the unconstrained and constrained image spectra $I(\vec{u})$ and $I_s(\vec{u})$ are uncorrelated. The sum of the variances in the real and imaginary parts of the two spectra, on the other hand, are approximately related by the convolution

$$N_s(\vec{u}) = \int d^2u_1 N(\vec{u} - \vec{u}_1)S(\vec{u}_1)e^{-(1/2)D_\phi(\lambda f u_1)}, \tag{5}$$

where $N(\vec{u}) = \langle |I(\vec{u}) - \langle I(\vec{u})\rangle|^2\rangle$ is that variance sum in the unconstrained image, N_s is the corresponding sum in the constrained image. $D_\phi(\lambda f u_1)$ is the structure function of the uncompensated phase errors at wavelength λ of the imaging light, and f is the focal length of the telescope. Note from Eq. (2) that

$$N(\vec{u}) = |O(\vec{u})|^2\sigma_H^2(\vec{u}). \tag{6}$$

By expanding the noise $N(\vec{u} - \vec{u}_1)$ as a power series in $\vec{u}_1$, a procedure that is justified for sufficiently narrow $S(\vec{u}_1)$ (equivalently, for sufficiently extended support size), we can express the change in the noise power spectrum under the application of support, i.e., $\Delta N(\vec{u}) \equiv N_s(\vec{u}) - N(\vec{u})$, to second order as

$$\Delta N(\vec{u}) = \vec{B} \cdot \vec{\nabla} N(\vec{u}) + \frac{1}{2}\hat{\mathbf{D}} : \vec{\nabla}\vec{\nabla} N(\vec{u}). \tag{7}$$

The vector $\vec{B}$ and the second-rank tensor $\mathbf{D}$, given by

$$\vec{B} = -\int d^2u S(\vec{u})e^{-(1/2)D_\phi(\lambda f u)}\vec{u} \tag{8a}$$

and

$$\hat{\mathbf{D}} = \int d^2u S(\vec{u})e^{-(1/2)D_\phi(\lambda f u)}\vec{u}\vec{u}, \tag{8b}$$

describe a decomposition of the noise transport process into a drift and a diffusion component, respectively.

It is worth noting that the tensor $\hat{\mathbf{D}}$ can always be diagonalized by a principal axis transformation, with the eigenvalues characterizing the two principal diffusion constants of the problem. The drift vector $\vec{B}$ represents, on the other hand, a ballistic transport of noise in the spatial frequency plane and is driven by the noise gradients in that plane. Note that the drift vector (8a) vanishes for any inversion-symmetric support, e.g., for a circular centered support, in which case the noise transport is purely diffusive. A more detailed discussion of the properties of noise transport and how that can be used to reduce the noise inside the support via the algorithm of Ref. 1 is presented in Ref. 2. We turn now to an analysis of the information dynamics that the preceding drift-diffusion dynamics of noise induces.

---

* Throughout the paper we mean, by the phrase "spectrum," the spatial frequency dependence of the variable being considered.

III. Information Dynamics under Support Constraint

Consider the mutual information $\mathcal{I}(\vec{u})$, which is the information successfully transmitted by an imaging system, at a given location $\vec{u}$ in the spatial frequency plane:

$$\begin{aligned}\mathcal{I}(\vec{u}) &= \mathcal{H}_O(\vec{u}) - \mathcal{H}_{O|I}(\vec{u}) \\ &= \mathcal{H}_I(\vec{u}) - \mathcal{H}_{I|O}(\vec{u}).\end{aligned} \tag{9}$$

The quantities $\mathcal{H}_O$, $\mathcal{H}_I$, $\mathcal{H}_{O|I}$, and $\mathcal{H}_{I|O}$ are the unconditional and conditional information entropies at the spatial frequency $\vec{u}$ in the statistical object and image ensembles, and are defined by the relations,

$$\mathcal{H}_O(\vec{u}) = -\int d^2O(\vec{u}) P_O(O(\vec{u})) \log P_O(O(\vec{u})); \tag{10a}$$

$$\mathcal{H}_{O|I}(\vec{u}) = -\int d^2O(\vec{u}) d^2I(\vec{u}) P(O(\vec{u}), I(\vec{u})) \log P_{O|I}(O(\vec{u})|I(\vec{u})); \tag{10b}$$

and the relations obtained from these by interchanging I and O. The quantity P_O is the (unconditional) pdf for the object spectral amplitude $O(\vec{u})$: $P_{O|I}$ is the conditional pdf for the object spectral amplitude $O(\vec{u})$, conditioned on the premise that a particular image spectral amplitude value $I(\vec{u})$ has been detected; and $P(O(\vec{u}), I(\vec{u}))$ is the joint pdf. These unconditional and conditional pdf's may also be regarded as pre and post detection pdf's for the object spectrum $O(\vec{u})$.

Before proceeding further, we note that we are led by the spirit of Shannon's information theory to deal with object and image ensembles, rather than single objects or images. The information theoretic treatment will thus address a class of objects, and is naturally disposed toward providing design and other evaluational considerations for an imaging system that may use a particular algorithm on a variety of objects.

The two alternative expressions for the mutual information in Eq. (9) convey the input-output symmetry of the information measure. When the prior knowledge of the object support is used, the image spectrum changes from Eq. (2) to the convolution (4). Such a change obviously causes no change in the object pdf P_O and therefore in the object information $\mathcal{H}_O(\vec{u})$. Thus the change $\Delta\mathcal{I}(\vec{u})$ in the mutual information when the support constraint is applied may be written as

$$\begin{aligned}\Delta\mathcal{I}(\vec{u}) &= -\Delta\mathcal{H}_{O|I}(\vec{u}) \\ &= \Delta\mathcal{H}_I(\vec{u}) - \Delta\mathcal{H}_{I|O}(\vec{u}).\end{aligned} \tag{11}$$

Although formally the change of $\mathcal{I}$ is more simply related to the conditional information $\mathcal{H}_{O|I}$, the computation of the conditional pdf $P_{O|I}$ is generally harder than that of the "inverse" conditional pdf $P_{I|O}$ that goes into the evaluation of $\mathcal{H}_{I|O}$. Indeed, both for the unconstrained and constrained image spectra (2) and (4), fixing the object spectrum $O(\vec{u})$ leads to conditional variates I and I_s that are linearly dependent on the Gaussian variate $H(\vec{u})$, as given by its pdf (3), and thus are themselves Gaussian variates. Their pdf's are

$$P_{I|O}(I(\vec{u})|O(\vec{u})) = \frac{1}{\pi N(\vec{u})} e^{-|I(\vec{u}) - O(\vec{u})\langle H(\vec{u})\rangle|^2 / N(\vec{u})} \tag{12a}$$

and

$$P^{(s)}_{I|O}(I_s(\vec{u})|O(\vec{u})) = \frac{1}{\pi N_s(\vec{u})} e^{-|I(\vec{u})-\langle I_s(\vec{u})\rangle|^2/N_s(\vec{u})}. \tag{12b}$$

where the superscript (s) refers to the support-constrained case and the symbols $N(\vec{u})$ and $N_s(\vec{u})$ refer to the image noise power spectral density before and after the support constraint is applied, for a fixed object spectrum $O(\vec{u})$. They are the same as the $N(\vec{u})$ and $N_s(\vec{u})$ defined in Sec. II and related by Eqs. (5) and (6). The conditional mean values $\langle I(\vec{u})\rangle$ and $\langle I_s(\vec{u})\rangle$ are given by averaging Eqs. (2) and (4) over the fluctuations of H for fixed $O(\vec{u})$.

The pdf's for the image spectrum, namely $P_I(I(\vec{u}))$ and $P^{(s)}_I(I_s(\vec{u}))$, are given by first multiplying Eqs. (12) by the object-spectrum probability density functional (pdF) $P_O[O(\vec{u})]$ (rather than the pdf at a single spatial frequency since all spatial frequencies occur in the convolution (5)) and then integrating over the entire object ensemble. From the so constructed image-spectrum pdf, one then has to use the formula (10a), with O replaced by I, to compute both the unconstrained information $\mathcal{H}_I(\vec{u})$ and the constrained one $\mathcal{H}^{(s)}_I(\vec{u})$, and the change $\Delta\mathcal{H}_I(\vec{u})$.

We have outlined a rather ambitious procedure of computing the change in the mutual information when support constraint is applied. Analytically, this is a prohibitive, if not impossible, task. We therefore introduce a simplifying assumption, namely that the change in the unconditional image information, $\mathcal{H}_I(\vec{u})$, is negligible compared to the that in the conditional information $\mathcal{H}_{I|O}(\vec{u})$. The justification for this assumption is that the pdf for the image spectrum results from an averaging of the conditional image pdf $P_{I|O}(\vec{u})$ over the statistical variations in the object ensemble. Such a statistical averaging tends to undo the frequency mixing caused by the convolution (4), since it makes the typical image spectral distribution in the ensmeble of such distributions have a more uniform character in frequency.

With this assumption, the change in mutual information is approximately the negative of the change in the conditional information $\mathcal{H}_{I|O}$:

$$\Delta\mathcal{I}(\vec{u}) \approx -\Delta\mathcal{H}_{I|O}(\vec{u}). \tag{13}$$

The change in the latter can be analytically computed, given the Gaussian pdf's (12) for the conditionals. Since

$$\mathcal{H}_{I|O}(\vec{u}) = -\langle \log P_{I|O}(I(\vec{u})|O(\vec{u}))\rangle$$

and since

$$\langle |I(\vec{u}) - \langle I(\vec{u})\rangle|^2\rangle = N(\vec{u}),$$

(and similarly for the support-constrained image) it follows that

$$\Delta\mathcal{H}_{I|O}(\vec{u}) = \left\langle \log\left[\frac{N_s(\vec{u})}{N(u)}\right]\right\rangle. \tag{14}$$

where the averaging is over the statistical variations in the object ensemble.

By performing the series expansion (5), dividing it by $N(\vec{u})$, and noting that

$$\frac{1}{N}\vec{\nabla}N = \vec{\nabla}\log N, \quad \frac{1}{N}\vec{\nabla}\vec{\nabla}N = \vec{\nabla}\vec{\nabla}\log N + \vec{\nabla}\log N\vec{\nabla}\log N,$$

we can express the change $\Delta\mathcal{H}_{I|O}$ as

$$\begin{aligned}\Delta\mathcal{H}_{I|O}(\vec{u}) \approx & \vec{B}\cdot\vec{\nabla}\mathcal{H}_{I|O} + \frac{1}{2}\vec{\mathbf{D}} : \vec{\nabla}\vec{\nabla}\mathcal{H}_{I|O} \\ & + \frac{1}{2}\vec{\mathbf{D}} : \left\langle \vec{\nabla}\log N\vec{\nabla}\log N\right\rangle - \left\langle (\vec{B}\cdot\vec{\nabla}\log N)^2\right\rangle \end{aligned} \tag{15}$$

consistent to second order. As for the noise, the transport of conditional information proceeds by means of ballistic motion (drift) as well as diffusion in the spatial frequency plane. However, unlike the noise transport, even for an inversion-symmetric support for which $\vec{B}$ vanishes, the transport of information remains sensitive to the noise gradients via the third term on the right hand side of Eq. (15).

V. Concluding Remarks

We have explored in this paper the subject of transport of Shannon's information in an imaging system when the prior knowledge of object support is employed. Understanding this transport is critical to divising an algorithm by which one can imporve the information throughput of the imaging system.

A related but much deeper question concerns a quantitative measure of information that one can associate with the *a priori* knowledge modalities like support. It seems intuitively obvious that a reconstructed image cannot possess more information than contained in the imposed *a priori* constraints and the raw measured data used in the reconstruction:

$$I[R] \leq I[C] + I[M], \tag{16}$$

where $I[R]$, $I[C]$, and $I[M]$ are the information measures of reconstructed image, imposed constraints, and measured data, respectively. But what is $I[C]$? Even if one could define it, it will, in general, depend on both the object statistics and the algorithm used in the reconstruction. It is clear, however, that for a given algorithm, one could pose an optimization question: For what object set is the information $I[C]$ in the imposed constraints maximized? Considerations analogous to those that address the capacity of communication channels must be employed to answer this question.

This work was partially supported by the Air Force Office of Scientific Research under Contract No. F49620-97-1-0148.

References

1. D. Tyler and C. Matson. "Reduction of nonstationary noise in telescope imagery using a support constraint." Opt. Express **1**. 347 (1997).

2. S. Prasad. "Dynamics of turbulence-induced noise in image deconvolution with support constraint," submitted to J. Opt. Soc. Am. A.

3. C. Shannon. "A mathematical theory of communication." Bell Syst. Tech. J. **27**. 379-423 and 623-656 (1948).

4. P. Fellgett and E. Linfoot. "On the assessment of optical images." Philos. Trans. R. Soc. (London) **A247**. 369-407 (1955).

5. E. Jaynes. "Information theory and statistical mechanics." Phys. Rev. **106**, 620-630 (1957).

6. B. R. Frieden. "Restoring with maximum likelihood and maximum entropy." J. Opt. Soc. Am. **62**, 511 (1972).

7. B. R. Frieden. "Image restoration using a norm of maximum information." Opt. Engineering **19**, 290-296 (1980).